

Marta Andronache

UNE DOUBLE ÉDITION PAPIER ET ÉLECTRONIQUE DE LA PREMIÈRE TRADUCTION EN FRANÇAIS DE LA CITÉ DE DIEU (1375)¹

Résumé : Offrir, dans le cadre d'un programme de recherche européen centré sur l'histoire de la langue, un glossaire des termes politiques et théologiques ainsi qu'une édition critique de la première traduction en français de la Cité de Dieu de saint Augustin, tel est l'objectif des travaux ici présentés. Si cette recherche – fondée sur un double support, papier et électronique – est appelée à constituer un outil précieux pour les linguistes, elle concerne aussi directement les historiens et les philosophes. Donner ainsi accès à la traduction de Raoul de Presles (1371-1375), c'est au-delà de l'expérience très riche d'une recherche assistée par l'informatique, faire rayonner un texte fondamental de la culture européenne pour répondre aux attentes d'une large communauté scientifique.

1. INTRODUCTION

Dans cet article, nous souhaitons présenter le programme de recherche « Histoire du vocabulaire politique français » qui se propose de donner un glossaire de termes politiques et théologiques et une édition critique de la première traduction en français de la *Cité de Dieu* de saint Augustin. Il s'agit d'un projet européen conçu en deux parties, un volet linguistique et un volet informatique, qui consistent à donner à la fois une édition critique sur support papier ainsi qu'une version électronique, et un glossaire de termes politiques.

Nous voudrions retracer le chemin de cette entreprise scientifique qui se met au service des linguistes, mais qui intéresse également des spécialistes d'autres domaines, comme les historiens ou les philosophes qui pourront s'en servir pour une large série d'études scientifiques à partir du texte de la *Cité de Dieu*.

Notre but est de montrer le travail mené sur l'informatisation des documents anciens et sur l'interface informatique/linguistique du projet. Dans ce cadre, nous découvrirons les étapes de la constitution d'une bibliothèque numérique autour de l'édition critique moderne de la *Cité de Dieu*, ensuite nous aborderons la question d'une double édition, papier et numérique, l'encodage XML et la conversion automatique en PDF, pour finir avec quelques aspects sur la constitution des glossaires qui accompagneront notre édition critique de texte et que nous allons constituer à partir d'un procédé de lemmatisation automatique.

¹ Mes remerciements les plus chaleureux s'adressent à Olivier Bertrand (ATILF-CNRS/Nancy Université) et à Béatrice Stumpf (ATILF-CNRS/Nancy Université) qui m'ont généreusement fait profiter de leurs réflexions stimulantes dans le cadre du projet ERC « Histoire du vocabulaire politique français » et qui ont bien voulu assurer une relecture critique du manuscrit, ainsi qu'à l'équipe d'informaticiens de notre projet : Bertrand Gaiffe, Jessika Pérignon, Etienne Petitjean, Gilles Souvay (ATILF-CNRS/Nancy Université).

2. « HISTOIRE DU VOCABULAIRE POLITIQUE FRANÇAIS », UN PROGRAMME DE RECHERCHE STARTING GRANTS ERC

« Histoire du lexique politique français » est un programme Starting Grants ERC lancé en 2008 et dirigé par Olivier Bertrand (CNRS, ATILF, Nancy). Ce programme s'étend sur cinq ans et se déroule au laboratoire CNRS « Analyse et traitement informatique de la langue française » (ATILF) à Nancy.

Qu'est-ce qu'un projet ERC ? C'est un programme qui a déjà été sélectionné dans le domaine des sciences humaines et sociales par Le Conseil européen de la recherche, (European Research Council - ERC)². Le public ciblé par les Starting Grants est formé de chercheurs en début de carrière en train de monter ou de consolider une équipe indépendante et qui souhaitent mener un projet de recherche exploratoire dans n'importe quel domaine de la science et de la technologie, dans un laboratoire situé dans l'Union européenne ou un État associé. ERC vise à stimuler l'excellence scientifique en Europe en soutenant les meilleurs chercheurs, dans tous les domaines par un grand financement qui est accordé pour une période de cinq ans. Les subventions permettent au chercheur de développer effectivement un projet, en assurant son salaire et le financement entier de son projet, y compris les ressources humaines nécessaires.

Le programme « Histoire du vocabulaire politique français », est justement un programme ERC de ce type. Ce programme vise à établir une édition critique de la première traduction en français de la *Cité de Dieu* de saint Augustin. Mais, à travers cette édition moderne de texte, le programme permet d'aller beaucoup plus loin qu'une simple édition de texte puisque finalement cet ouvrage scientifique ouvre à un public bien plus large l'accès à ce texte fondamental de la culture européenne.

Les principales thématiques de recherche de notre programme sont représentées par des études lexicologiques et diachroniques du français, qui prennent en compte la philologie, la codicologie et la paléographie, ainsi qu'à des études sur l'histoire des idées politiques.

L'équipe qui travaille dans le cadre de ce programme a été constituée par le directeur du programme et elle est formée de neuf transcripteurs dont trois travaillent à l'ATILF, les autres travaillent dans d'autres laboratoires et universités de France, d'Italie et du Canada. Les rédacteurs de l'édition critique sont soutenus par une équipe informatique et TEI, une équipe de documentation, un réseau de chercheurs et de collaborateurs externes, ainsi qu'un comité scientifique.

3. L'ÉDITION CRITIQUE

Face à un texte comme la *Cité de Dieu* plutôt bien connu du monde scientifique par des éditions et traductions modernes de ce texte latin qui remonte à 426, on peut se poser la question suivante : pourquoi est-il important de donner une édition critique à la première traduction en français de *La Cité de Dieu* de saint Augustin dont la rédaction s'est achevée en 1375 ?

Les traductions en moyen français des textes de science politique (grecs et latins) de l'antiquité ne sont que partiellement éditées. C'est par les traductions en moyen français des textes

² Le programme ERC « Starting Grant 2007 » dont est issu celui qui fait l'objet du présent article est l'un des 300 projets lauréats sur 9200 candidats européens, toutes matières confondues.

savants grecs et surtout latins que se constitue une grande partie du lexique que nous utilisons en français aujourd'hui. C'est le cas d'une partie du lexique de la science politique qui naît en France aux 14^e et 15^e siècles, précisément à la faveur des traductions des œuvres de l'antiquité. Il s'agit de traduire des concepts qui n'existent pas en français et qui pourtant ont une existence avérée en latin ou en grec. Il est bien connu que le texte de *La Cité de Dieu* est un ouvrage fondamental dans l'histoire des idées en Occident (Bertrand 2003 et 2004, Duval 2007). La richesse du texte ne se limite pas à la seule traduction du texte latin en français, puisqu'il bénéficie d'importantes gloses du *translateur* médiéval, ce qui offre un champ d'investigation scientifique majeur en matière de lexicologie historique, histoire, philosophie. Grâce à une édition moderne de la première traduction en français (1371-1375) de *La Cité de Dieu* de saint Augustin par Raoul de Presles, un public assez large aura la possibilité de mieux connaître l'histoire du lexique politique et théologique en français.

Le but du projet est l'étude lexicologique du vocabulaire politique et théologique à partir d'une édition critique de la première traduction en français de *La Cité de Dieu* par Raoul de Presles, rédigée entre 1371 et 1375 et dont les premiers manuscrits apparaissent vraisemblablement entre 1376 et 1380. Les avantages d'une double édition sont évidents : la version papier est figée, mais elle permet une certaine autonomie de la lecture, la possibilité de la consulter à tout moment, de prendre des notes et elle répond à certaines habitudes des lecteurs. L'édition numérique se prête à un tout autre type d'utilisation du texte : pour extraire par diverses requêtes des informations qui constitueront le départ d'études interdisciplinaires autour de ce texte majeur du Moyen Âge.

Cette édition aura pour manuscrit de base l'exemplaire de la bibliothèque royale de Charles V (1364-1380) conservé en deux volumes sous la double cote FR 22912 et FR 22913 de la BnF, département des manuscrits occidentaux. Au moment où l'inventaire de la bibliothèque de Charles V a été établi, en 1380, ce manuscrit y était répertorié. Il s'agit d'un manuscrit dont la date d'écriture est très proche de la date de composition de l'œuvre par Raoul de Presles (Laborde 1909). Par conséquent, même si ce n'est pas le manuscrit original de la première traduction, nous sommes en présence d'un manuscrit royal d'une très bonne qualité et qui est très haut placé dans le stemma ou dans la famille des manuscrits de la première traduction de *La Cité de Dieu* en français. À côté de ce manuscrit, nous disposons encore d'autres manuscrits qui datent de la même période, depuis la fin de la rédaction de l'ouvrage par Raoul de Presles (environ 1375) jusqu'à la fin du 14^e siècle. C'est le cas du manuscrit de la bibliothèque du Château à Chantilly daté depuis la fin du 14^e siècle, de trois des manuscrits de la BnF datant respectivement de 1380, de *ca.* 1385 et de 1390, d'un des manuscrits de Bruxelles de *ca.* 1390, et d'un autre manuscrit de la Bibliothèque royale de La Haye daté de *ca.* 1390, pour ne citer que quelques-uns des neuf manuscrits de contrôle choisis pour établir l'édition critique de la première traduction en français de la *Cité de Dieu*. Aujourd'hui nous avons pu identifier 60 manuscrits de la *Cité de Dieu* qui couvrent la période qui va d'environ 1380 jusqu'au 15^e siècle, avec l'apparition d'un incunable de 1486 (Abbeville) et une édition datant de 1531.

4. DE LA TRADITION MANUSCRITE VERS UNE BIBLIOTHÈQUE NUMÉRIQUE DE LA CITÉ DE DIEU

Devant la richesse de la tradition manuscrite du texte, est apparue la nécessité de constituer une bibliothèque numérique comme une étape de travail préparatoire de notre édition critique. Le

but est de réunir toutes les reproductions des manuscrits de la *Cité de Dieu* qui nous ont été transmis et qui sont conservés dans les bibliothèques du monde entier. De cette manière, ces précieux documents deviendront facilement accessibles à tous les chercheurs impliqués dans le projet.

Pour établir la liste des manuscrits qui portent le texte, nous avons au cours d'une première étape utilisé l'ouvrage de Laborde (Laborde 1909) que nous avons complété entre temps. Seuls les manuscrits complets et datés jusqu'au 15^e siècle ont été pris en compte. Dans une deuxième étape, nous avons acquis les microformes ou CD-Rom des manuscrits les plus anciens pour constituer le fonds ERC à l'ATILF. Par la suite, ces documents sont scannés et enregistré sous une forme numérique sous un espace web où ils sont accessibles à tous les membres de l'équipe.

Aujourd'hui, le fonds ERC compte déjà 32 copies de manuscrits sous forme de CD-Rom, microfilms ou microfiches. Des négociations sont en cours actuellement auprès d'autres bibliothèques pour enrichir notre fonds. En raison d'une tradition manuscrite exceptionnelle, la complémentation de la bibliothèque numérique représente une masse énorme de travail.

Un des aspects importants de notre bibliothèque numérique est qu'elle est déjà fonctionnelle : elle est accessible sur un espace ERC consultable à distance par tous les membres de l'équipe munis d'un *login* personnalisé et d'un mot de passe.

On peut se demander à quoi sert tout ce travail de numérisation des anciens manuscrits et la création de cette bibliothèque virtuelle. On relève au moins deux aspects très pratiques : premièrement, bénéficier dans un fonds unique de tous les manuscrits de la *Cité de Dieu* réunis, en provenance de toutes les bibliothèques du monde. Deuxièmement, permettre aux membres de notre équipe de confronter rapidement pour leurs transcriptions de texte, par exemple, ou pour l'appareil critique de l'édition, la consultation de ces documents et le manuscrit de base avec d'autres manuscrits de contrôle. Ainsi, une équipe séparée par des milliers de kilomètres, gagne dans l'harmonie et dans la cohérence de la méthode de travail. De cette manière, la bibliothèque numérique de *La Cité de Dieu* constitue une richesse en soi, mais elle présente aussi une haute fonctionnalité dans le travail d'édition de texte.

5. LA PRÉPARATION DE LA DOUBLE ÉDITION CRITIQUE DE TEXTE

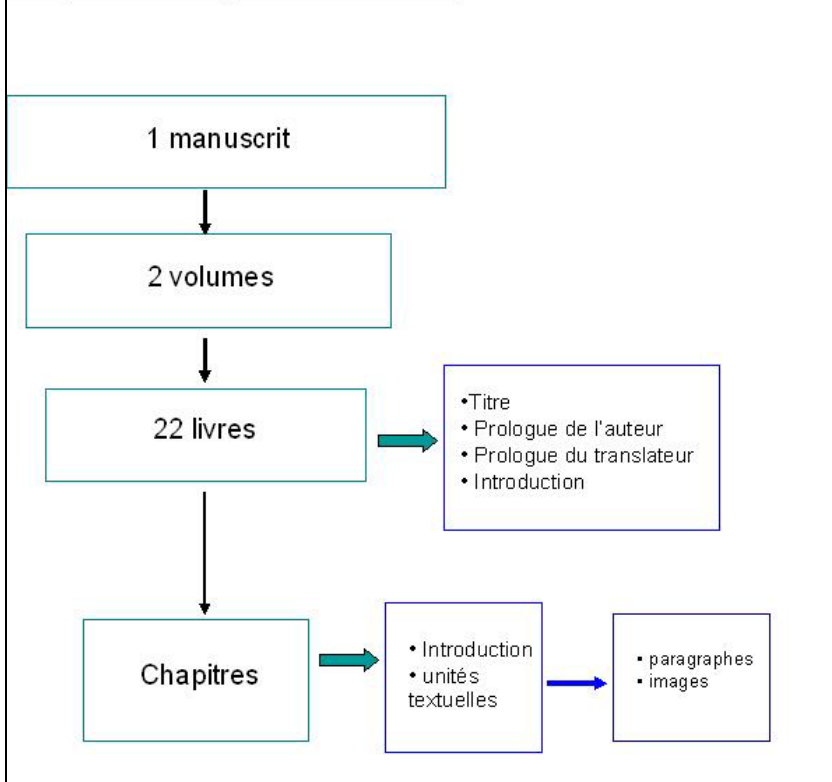
Un autre aspect de l'informatisation de notre édition de texte concerne l'encodage XML et la conversion automatique en PDF.

5.1. Le cahier de charges informatiques

La réalisation du schéma XML a représenté un des grands chantiers de travail de la facette informatique de notre projet ERC.

Pour établir le dialogue entre les linguistes et l'équipe informatique, nous avons commencé par établir un cahier de charges dont l'objet était de présenter les données matérielles du manuscrit aux informaticiens. Cela a été réalisé par la description concrète du manuscrit. Pour donner un exemple, nous avons schématisé la disposition et la segmentation du texte de notre manuscrit de base en deux volumes, 22 livres, qui contiennent chacun plusieurs dizaines de chapitres et qui, à leur tour, sont constitués du texte de la traduction suivi d'un commentaire du traducteur qu'il appelle « exposition ». De surcroît, le 1^{er} volume présente des éléments spécifiques : un incipit qui comprend une liste des autorités, une dédicace et un prologue.

II. Disposition et segmentation du texte :

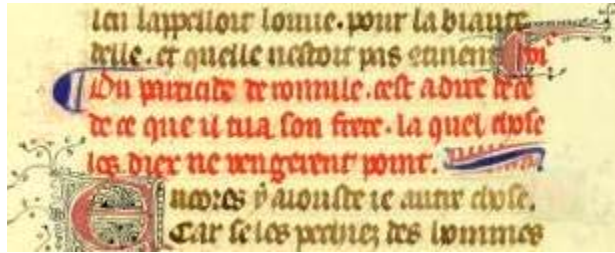


Au-delà de ce type d'informations de structuration du texte de notre manuscrit, le Cahier des charges informatiques comprend aussi des informations concernant les éléments à encoder. Mais pour décrire ce qu'on voudrait encoder, on a dû préalablement se poser la question suivante : que convient-il d'encoder dans le cas spécifique de notre projet ? Pour résoudre ce problème de départ, notre idée a été que l'encodage permet de faire ressortir toutes les informations pour déterminer par la suite leur destination en fonction de l'utilisation que l'on entendra faire de ce texte encodé : édition papier, édition numérique, glossaire du vocabulaire politique, théologique, général ou recherches pour des études scientifiques.

De manière générale, nous avons décidé de baliser des éléments qui correspondent à une catégorie spécifique préalablement déterminée dans notre Cahier des charges informatiques.

Il s'agit tout d'abord des marques spécifiques du texte :

- la séparation en livres ;
- la séparation en chapitres et les titres :



(BNF, fr. 229112)

- les paragraphes du type traduction et les paragraphes du type glose appelés « expositions » par le traducteur ;
- les fins de colonnes et les fins de folios ;
- les pieds de mouche comme marque de séparation du texte par le copiste :



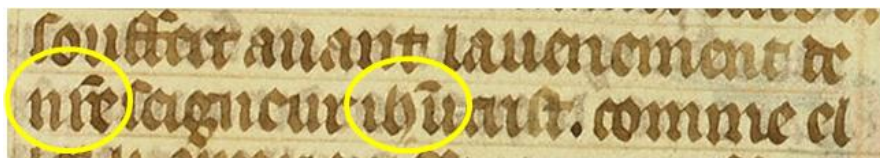
(BNF, fr. 229112)

- les lettrines :



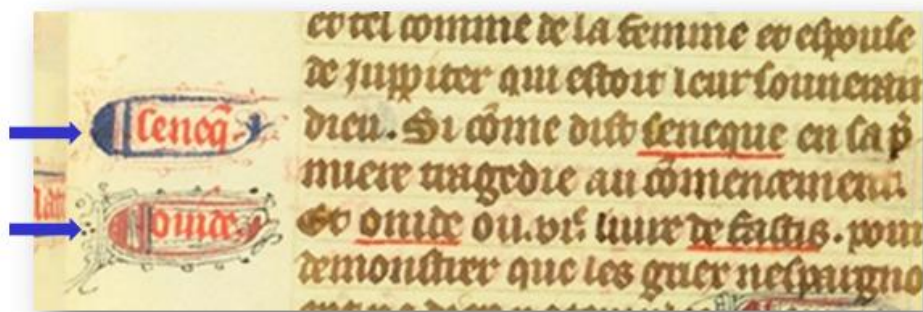
(BNF, fr. 229112)

- la résolution des abréviations :



(BNF, fr. 229112)

- les additions marginales présentes dans le manuscrit :



(BNF, fr. 229112)

Ensuite nous avons déterminé d'autres catégories à encoder comme, par exemple, les lexèmes, destinés à entrer dans les glossaires et dans les index, ou bien les citations et les sources invoquées par le traducteur.

Les corrections constituent un autre type d'information encodée. Par exemple, en cas de texte manifestement fautif ou lacunaire, une correction au manuscrit de base s'impose et l'information est encodée par une balise <corr> de type "fautif", balisage repérable lors d'une recherche ultérieure :

969	<g ref="#pdm"/> De l'<term type="politique" key="iniquité"
970	>iniquité</term> de la <term type="politique" key="bataille"
971	>bataille</term> que les <index indexName="ethnique"><term
972	key="Romain">Romains</term></index> firent aus <corr
973	type="fautif" source="#C1 #P17 #P11">Albains</corr>
974	<note type="critique">Ms P1: Albams; Mss C1, P17, P11 : Albains;
975	P31: Albams</note> et de la <term type="politique"
976	key="victoire">victoire</term> prise pas eulz par <term
977	type="général" key="convoitise">convoitise</term> de <term
978	type="politique" key="seigneurir">seigneurir</term>.

Dans l'édition papier, cette erreur sera signalée au moyen de crochets qui se fera de manière automatique, comme nous pouvons le constater ci-dessous :

·XIII·	De l'iniquité de la bataille que les Romains firent aus [Albains] ^[1] et de la victoire prise pas eulz par convoitise de seigneurir.
--------	---

De la même manière, nous avons encodé plusieurs types de corrections : lorsqu'une leçon est rejetée, à la suite d'une inadvertance du copiste (<corr type="remplacement" source="#A #B">frere</corr>) ; quand il y a une erreur manifeste de transcription de la part du copiste, comme par exemple quand le texte original donne « prece livre dent », la correction « livre precedent » est faite (<corr type="fautif">livre precedent</corr>).

Les corrections sont suivies en général par une note critique qui permet de savoir ce qui se trouvait dans le manuscrit avant la correction et elles sont toujours accompagnées de l'attribut

« source » qui donne explicitement les manuscrits de contrôle qui nous ont permis de faire les changements.

Il est évident que nous avons dû opérer un certain nombre des choix dictés par la nature de notre texte qui est en prose et qui est très ample. Par exemple nous avons décidé de ne pas marquer les fins de ligne, mais seulement les fins de colonnes et de folios, mais par contre, il est prévu un lien vers l'image de chaque folio pour donner accès à plus d'informations, informations qu'un chercheur voudrait récupérer par la suite (uniquement dans le cadre de la version numérique évidemment).

En fin de compte, l'équipe s'est aussi interrogée sur le balisage, puisque, d'une part, il s'agit d'un processus qui ralentit de manière évidente la transcription et d'autre part, les éditions d'ancien texte se font souvent sous Word. Les avantages sont en définitive évidents puisque les balises ont une double finalité : - à l'encodage, elles garantissent une harmonisation interne dans la manière de transcrire (et il convient que l'ensemble de l'édition soit homogène de ce point de vue) et dans la méthode de travail choisie ; - au décodage, les balises permettent l'interrogation du texte selon des multiples critères, par exemple des recherches sur les différents types de lexique et pour tout autre type d'études.

5.2. Encodage XML

La réalisation du schéma XML a représenté un des grands chantiers de travail correspondant à la facette informatique de notre projet ERC.

Sur la base de ces informations, l'équipe informatique a constitué le schéma XML-TEI que nous avons discuté, testé, amélioré et rediscuté tout au long des réunions de notre équipe avec les informaticiens du laboratoire.

Du point de vue linguistique, l'important était d'avoir un texte structuré et hiérarchisé selon un protocole prédéfini mais le défi restait d'encoder les informations critiques de notre texte par des balises, ce qui n'est ni le plus rapide, ni le plus facile à faire!

Une fois le schéma XML prêt, nous avons pu commencer de transcrire notre texte sous ce format, même si des modifications et des améliorations peuvent encore se faire.

Le résultat correspond à un document que nous pouvons ouvrir sous ce mode ayant l'avantage de bien relever la structuration du texte, mais qui peut sembler, à première vue, illisible, même rébarbatif, comme nous pouvons le voir dans cet exemple avec la transcription en mode texte (BNF, fr. 229112 : III, 1, 97ab) :

1165 </div>
 1166 <div type="chapitre" n="1">
 1167 <head><g><#pdm?><choice><abbr>C.</abbr><abbr>Chapitre</abbr></choice>
 1168 <num></num> Des aversités que seulement les malvaiz craignent et lesquels
 1169 le monde à toujours souffers <w>ta-g>n</g><w> comme il a aoure les ydoles </head>
 1170 <div type="traduction">
 1171 <p><cb n="c"/>
 1172 <w>
 1173 <c type="letrine" rend="cap4">C.</c><w> cuide que j'aye asses dit
 1174 des mauvaises <term type="theologique" key="meurs">meurs</term> de
 1175 courage qui sus toutes choses sont a eschever, et que les faux <term
 1176 type="theologique" key="dieux">dieux</term> ne tindrent compte de
 1177 secourir au <term type="politique" key="peuple">peuple</term> qui les
 1178 aouroit, afin <w>q-g>ue</g><w> ilz ne feissent ces maux mais, qui plus
 1179 est, ont fait que il ont esté <w>g-g>n</g>strains</w> à les faire. <g
 1180 re="pdm?> Mais il semble que il est à dire des mauz lesquelz
 1181 yceux seulement ne veulent souff<cb n="d"/>rir, si comme
 1182 <w>so-g>n</g><w> fain, maladie, batailles, pillerie, chetiveté,
 1183 occision, et autres, s'aucuns en ya samblables dont nous avons fait
 1184 memoire ou premier livre. Car ces choses les mauvais seulement deputent
 1185 à mal jasoit ce que elles ne les facent pas mauvais, ne il n'ont pas
 1186 honte d'estre mauvais entre les <w>b<g>le</g>n<w> de <index
 1187 indexName="anthroponyme">
 1188 <term key="Fortune">Fortune</term>
 1189 </index> que il loent et desirrent à avoir. Et se <term type="général"
 1190 key="courroucer">courroucer</term> plus se il ont mauvaise ville
 1191 que se il <w>o-g>n</g><w>
 1192 <cb n="96?><cb n="3?> mauvaise vie, c'est à dire que il soyent plains
 1193 de mauvaises <term type="theologique" key="meur">meurs</term> aussi
 1194 comme se ce fust le tres grant bien de l'omme avoir tous biens et
 1195 <w>n<g>n</g><w> soi avoir. <g re="pdm?> Ne leurs <term
 1196 type="theologique" key="dieux">dieux</term> ne resisterent à ce que
 1197 ces maux que il <term type="général" key="ressoigner">
 1198 >ressongnient</term> seulement ne leur avenissent quant il les aouroient
 1199 si franchement et liberaument. Car, comme avant l' <term
 1200 type="theologique" key="avènement">avènement</term> de
 1201 <w>n<g>ost</g>re</w> seigneur <index indexName="anthroponyme">
 1202 <term key="Jesus_Christ">
 1203 <w>Jh<g>es</g>u Crist</w>
 1204 </term>
 1205 </index> par divers lieux et par divers temps l'umain <term
 1206 type="général" key="lignage">lignage</term> fust <term type="général"
 1207 key="tourmenter">tourmenter</term> de tant de batailles sans nombre
 1208 et de telz qui ne <w>se<g>m</g>blent</w> pas à croire quelx autres <term
 1209 type="theologique" key="dieux">dieux</term> aouroit le monde que
 1210 ceuk que l'en aoure à present, excepté par aventure <num></num>
 1211 <term type="politique" key="peuple">peuple</term>
 1212 <index indexName="ethnique">
 1213 <term key="Hebreu">hebreu</term>

On peut simplifier l'aperçu et lire le même texte dans un mode où les balises s'affichent sous forme d'étiquettes, ce qui nous permet de bien suivre notre processus de balisage, mais qui donne un texte qui reste encore difficile à lire pour une relecture, par exemple (Transcription en mode auteur avec balises, BnF, fr. 229112 : III, 1, 97ab) :

g C. Chapitre -I- Des averstés que seulement les malvaiz craignent et lesquels le monde a tousjours souffers tant comme il a aoure les ydoles

cb C euide que j'aye asses dit des mauuaises meurs de courage qui sus toutes choses sont a eschever, et que les faux diex ne tindrent compte de secourir au peuple qui les aouroit, afin q'ilz ne feissent ces mauiz mais, qui plus est, ont fait que il ont esté costraints à les faire. Mais il semble que il est à dire des mauiz lesquels yeuix seulement ne veulent souffrir, si comme faim, maladie, batailles, pillerie, chetivété, occision, et autres, s'aucuns en ya samblables dont nous avons fait memoire ou premier livre. Car ces choses les mauuaises seulement deputent à mal jasoit ce que elles ne les facent pas mauuaises, ne il n'ont pas honte d'estre mauuaises entre les pb. Cest assavoir à la cité de Rome proprement et des terres quelconques qui li sont conjointes par société ou par compagnie ou par

TEI text body div div div div p note

g C. Chapitre -I- Des averstés que seulement les malvaiz craignent et lesquels le monde a tousjours souffers tant comme il a aoure les ydoles

cb C euide que j'aye asses dit des mauuaises meurs de courage qui sus toutes choses sont a eschever, et que les faux diex ne tindrent compte de secourir au peuple qui les aouroit, afin q'ilz ne feissent ces mauiz mais, qui plus est, ont fait que il ont esté costraints à les faire. Mais il semble que il est à dire des mauiz lesquels yeuix seulement ne veulent souffrir, si comme faim, maladie, batailles, pillerie, chetivété, occision, et autres, s'aucuns en ya samblables dont nous avons fait memoire ou premier livre. Car ces choses les mauuaises seulement deputent à mal jasoit ce que elles ne les facent pas mauuaises, ne il n'ont pas honte d'estre mauuaises entre les pb. Cest assavoir à la cité de Rome proprement et des terres quelconques qui li sont conjointes par société ou par compagnie ou par

Et, par un simple clic, on peut afficher et écrire dans un mode absolument lisible qui peut cacher les balises et présenter un texte propre et facile à lire ou relire :

TEI text body div div div div p note

g C. Chapitre -I- Des averstés que seulement les malvaiz craignent et lesquels le monde a tousjours souffers tant comme il a aoure les ydoles

cb C euide que j'aye asses dit des mauuaises meurs de courage qui sus toutes choses sont a eschever, et que les faux diex ne tindrent compte de secourir au peuple qui les aouroit, afin q'ilz ne feissent ces mauiz mais, qui plus est, ont fait que il ont esté costraints à les faire. Mais il semble que il est à dire des mauiz lesquels yeuix seulement ne veulent souffrir, si comme faim, maladie, batailles, pillerie, chetivété, occision, et autres, s'aucuns en ya samblables dont nous avons fait memoire ou premier livre. Car ces choses les mauuaises seulement deputent à mal jasoit ce que elles ne les facent pas mauuaises, ne il n'ont pas honte d'estre mauuaises entre les pb. Cest assavoir à la cité de Rome proprement et des terres quelconques qui li sont conjointes par société ou par compagnie ou par

c Exposition sus ce chapitre

Translatour En ce tiers livre monseigneur saint Augustin entent à traiter des mauz de painne ausquelz sont soubzmas ceulz qui aourent leurs diex. Et est son entencion de monstrier que leurs diex que ilz aourent ne les gardent point de ces mauz. Et que pour ceste cause, ne a ceste entencion, l'en ne les doit point aouer. Et pource, en cest premier chapitre, il declare son entencion et comment il veult proceder en cest livre et, par especial, quant à la matiere de laquelle il veult traiter. C'est assavoir des mauz de painne et non pas de toutes manieres de gens, mais des Rommains et de ceulz qui ont compaignie avec eulz, si comme les Italiens et les autres nations que ilz mistrent en subjection depuis le temps que ilz les subjuguèrent. Après, quant il parle et fait exception du peuple hebreu et d'aucuns autres hors ycelli peuple qui ont esté reservez, il semble qu'il se puest entendre pour une maniere de gent qui sont appelez bragmann ou gignosphistes ou pour les vechabites. De ces bragmanns parle monseigneur saint Gerome Ierominus ou prologue de la Bible. Et dit que ilz sont en Judé et que ce sont une gent qui vacquent à apprendre science et apprendre le cours des estoilles. A iceulz vult fere guerre Alexandre mais il appelle Didimus qui estoit leur maistre et leur évesque l'appaissa par certaines lectures que il li envoya. Esquelles il lui escript toute leur vie si comme il le treuve es fais d'Alexandre. Lesquelles epistres tu pourras veoir par Vincent in Speculo Historali Vincent es temps de Alexandre. Et quant est des gignosphistes cb samblablement en parle monseigneur saint Gerome Ierominus ou dit premier

Aujourd'hui nous travaillons avec un schéma RelaxNG (RNG) que nous mettons à jour automatiquement sur une adresse <http>. L'équipe informatique travaille à l'amélioration de cet outil qui permet un certain nombre de fonctionnalités telles que la génération automatique d'un nouveau schéma RNG, la mise à jour et l'harmonisation des fichiers XML par rapport aux travaux des autres transcribers, l'enrichissement automatique des listes de normalisation pour le balisage des index et des glossaires.

5.3. Conversion PDF

Une partie du travail informatique consiste à mettre au point la possibilité d'une conversion PDF adaptée à notre projet. Ce travail a une double utilité : pour l'édition électronique il est adapté aux relectures du texte transcrit ; pour l'édition papier, la conversion du XML en PDF nous permet d'obtenir notre texte sous un format de présentation du texte à l'éditeur.

Voici un exemple de conversion en PDF du texte transcrit en XML dont nous avons vu quelques exemples ci-dessus :

Exposicion sus ce chapitre

Translateur

En ce tiers livre monseigneur saint Augustin entent à traitier des maux de painne ausquelz sont soubzmis ceulz qui aourent leurs diex. Et est son entencion de monstrier que leurs diex que ilz aourent ne les gardent point de ces maux. Et que pour ceste cause, ne a ceste entencion, l'en ne les doit point aouer. Et pource, en cest premier chapitre, il declere son entencion et comment il veult proceder en cest livre et, par especial, quant à la matiere de laquelle il veult traitier. C'est assavoir des maux de painne et non pas de toutes manieres de gens, mais des Rommains et de ceulz qui orent compaignie avec eulz, si comme les Ytaliens et les autres nations que ilz mistrent en subjection depuis le temps que ilz les subjuguèrent.

Après, quant il parle et fait exception du pueplehebreu et d'aucuns autres hors ycelli pueple qui ont este reservez, il semble qu'il se pueist entendre pour une maniere de gent qui sont appelez bragmani ou gignosphiste ou pour les vechabites.

Ierominus

De ces bragmanis parle monseigneur saint Gerosme ou prologue de la *Bible*. Et dit que ilz sont en Judé et que ce sont une gent qui vacquent à apprendre science et apprendre le cours des estoilles. A iceulx vould fere guerre Alixandre mais -I- appellé Didimus qui estoit leur maistre et leur evesque l'appaisa par certaines lectres que il li envoya. Esquelles il lui escript toute leur vie si comme il le treuve es fais d'Alixandre. Lesqueles epitres tu pourras veoir par Vincent in *Speculo Historiali* es temps de Alixandre.

Vincet

Ierominus

Et quant est des gignosphistes^[f^o95v^o] samblablement en parle monseigneur saint Gerosme ou dit premier prologue. Et dit que ilz sont en Ethiopie et sont gent de merveilleuse condicion. Car ilz sont touz nulz et sont à une table d'or sur le sablon où ilz ne s'estudient que à vraie contemplacion et procedent en humilité et poureté et habitent en petites maisons et en cavernes. Ne n'ont ne villes ne citez, ne nuyent à nulluy, ne ilz n'ont cure d'ameures et demeurent arrieres de leurs femmes et de leurs enfans qui nourrissent les bestes de quoy ilz doivent vivre. Yceulz trouva Alixandre en Ethiopie, et quant il les ot trouvez il fu moult esmerveillez de leur vie et de leur ordenance et la pris moult. Et leur dist que il demandassent ce que il vouldroyent et il leur dourroit. Lesquies li requisrent que il leur donnast immortalité et que c'estoit la chose que ilz desirroyent plus ou monde, et que de ses richesses ilz ne tenoyent compte. Lequel leur respondi que c'estoit chose impossible que il leur peust donner immortalité comme il feust homme mortel. Et il lui respondirent comment donques il faisoit tant de maulz par tout le monde, puis que il sentoit que il estoit homme mortel, si comme il se treuve ou *Livre des merveilles d'Orient*.

Bible

Et quant est des vechabites, il s'en treuve en la *Bible* ou ·XXXV^e chapitre de

Les difficultés rencontrées avec ce processus sont multiples, mais l'équipe informatique de notre projet a déjà résolu une grande partie des problèmes posés en utilisant le logiciel LaTeX, comme par exemple celui correspondant au calcul de l'alignement pour les notes en marge de page ou bien la mise en place des différents niveaux des notes de bas de pages et leur ancrage dans le texte.

Notre souci permanent a été d'offrir un texte lisible, qui ne soit pas noyé dans des appels de notes difficiles à suivre et qui soit muni d'un appareil critique pratique à manier. Ce processus est en cours de réalisation.

6. LA CONSTITUTION DU GLOSSAIRE

Notre édition de texte sera accompagnée d'un glossaire. Nous avons opté pour la lemmatisation du texte, mais nous avons gardé aussi un balisage pour certains mots, celui des noms propres par exemple, difficilement détectés comme noms propres par le lemmatiseur.

La lemmatisation nous permet de grouper les formes sous les lemmes correspondants et de bénéficier d'une série de fonctionnalités très utiles à la fois pour la construction automatique de glossaires et pour les recherches ultérieures d'une édition électronique de manuscrit (Souvay/Pierrel 2009). Ainsi, nous pouvons disposer de la nomenclature des lemmes du texte lemmatisé (*La Cité de Dieu*, Livre III) :

Dictionnaire du Moyen Français (1330-1500)

Glossaire : nomenclature des lemmes

■ Nomenclature livre3

Code	Lemme	Catégorie	Fréq. 1	Fréq. 2
[1067]	PACTE	subst.	1	1
[1068]	PAILLER1	verbe	1	1
[1069]	PAIN	subst.	1	1
[1070]	PAIR	adj.	1	109
[1071]	PAIRE	subst.	1	5
[1072]	PAITRE	verbe	1	1
[1073]	PAIX	subst.	2	5
[1074]	PALLIER	verbe	1	1
[1075]	PARD	subst.	1	1
[1076]	PAREIL	subst.	2	4
[1077]	PAREILLEMENT	adv.	1	1
[1078]	PARELLE	subst.	1	1
[1079]	PARENT	subst.	1	1
[1080]	PARER1	verbe	1	5
[1081]	PARIS	adj.	1	5
[1082]	PARJURE1	subst.	1	2
[1083]	PARJUREMENT	subst.	2	10
[1084]	PARJURER	verbe	5	7
[1085]	PARLANT	adj.	2	2
[1086]	PARLER	verbe	7	26

Nous pouvons aussi repérer les mots inconnus par le DMF :

DMF

ID	Mot	Catégorie	Texte	Page
4	Mitridaces	mot inconnu	esté tres bonne. XXII De l'edict que fist Mitridaces par lequel il commanda que touz les citoyens rommains	page 94
5	bragmani	mot inconnu	puest entendre pour une maniere de gent qui sont appellez bragmani ou gignosophiste ou pour les vechabites. De ces	page 95
6	gignosophiste	mot inconnu	pour une maniere de gent qui sont appellez bragmani ou gignosophiste ou pour les vechabites. De ces bragmanis parle	page 95
7	vechabites	mot inconnu	gent qui sont appellez bragmani ou gignosophiste ou pour les vechabites. De ces bragmanis parle monseigneur saint Gerosme ou	page 95
8	bragmanis	mot inconnu	bragmani ou gignosophiste ou pour les vechabites. De ces bragmanis parle monseigneur saint Gerosme ou prologue de la Bible	page 95
9	historiali	mot inconnu	Lesqueles epistres tu pourras veoir par Vincent in Speculo Historiali es temps de Alixandre. Et quant est des	page 95
10	gignosophistes	mot inconnu	Historiali es temps de Alixandre. Et quant est des gignosophistes samblablement en parle monseigneur saint Gerosme ou dit premier	page 95

48 réponses << Page 1 / 5 >> Actualiser Options ABC

Pour les mots inconnus, le lemmatiseur fait des propositions ou indique qu'il n'y a pas cette possibilité ; nous disposons aussi d'une fonctionnalité qui nous permet de retirer ou de valider des lemmes lorsqu'une ambiguïté a été levée.

Il est possible de repérer le nombre de mots sous la forme d'une liste avec un lien qui nous permet d'afficher aussi les contextes, les mots en ordre d'apparition ou par ordre alphabétique, d'éliminer du bruit par l'application ou la création de nouveaux filtres.

Pour la lemmatisation, nous travaillons en collaboration avec l'équipe du *Dictionnaire de Moyen français (DMF)* et le lemmatiseur pourra être adapté aux besoins du glossaire de notre édition de texte.

Ce travail est en cours, mais nous pouvons déjà apprécier l'utilité de cet outil qui permet au linguiste, une fois le travail de transcription fini, par exemple, de changer de perspective et de travailler comme lexicographe ce qui augmente la fiabilité du glossaire qui accompagnera notre édition critique de *La Cité de Dieu*.

7. CONCLUSION

En conclusion, aujourd'hui nous avons à notre disposition le schéma XML dans lequel nous pouvons transcrire le texte de notre manuscrit qui sera transformé en un document moderne. Malgré les contraintes liées au balisage, ce schéma présente d'importants avantages, puisque ce type de fichier permet de disposer d'une méthode harmonieuse de travail, de récupérer de l'information pour des recherches ultérieures, il permet le traitement informatique du texte, et, par un simple clic, nous fait bénéficier d'un texte propre et lisible tout à fait prêt à un double emploi, pour l'édition papier et pour l'édition électronique.

Par la méthodologie de travail rigoureuse que nous nous sommes imposés et par les outils informatiques performants dont nous disposons à l'heure actuelle dans le cadre du projet ERC « Histoire du lexique politique français », il nous sera possible, dans un délai raisonnable, de donner l'accès à la première traduction en français de la *Cité De Dieu* par Raoul de Presles (1371-1375), que ce soit sous la forme papier ou sous forme numérique. Notre espoir reste que ce texte fondamental de la culture européenne réponde à l'attente des linguistes et d'une large communauté scientifique.

Marta Andronache
ATILF (CNRS/Nancy Université)

BIBLIOGRAPHIE

- AUGUSTAIN, saint, *La Cité de Dieu*, volume I, ms BNF fr. 22912, Paris.
- BERTRAND, O. (2004). *Du vocabulaire religieux à la théorie politique en France au XIV^e siècle. Les néologismes chez les traducteurs de Charles V (1364-1380)*, Paris, Connaissances et Savoirs.
- BERTRAND, O. (2003). « Les néologismes politiques dans la première traduction française de *La Cité de Dieu* de saint Augustin 1375 », *The Theory and Practice of Translation in the Middle Ages, (The Medieval Translator. Traduire au Moyen Âge, 8)*, R. Voaden, R. Tixier, T. Sanchez Roura et J. R. Rytting (éd.), Turnhout : Brepols, 39-48.
- BOSSUAT, R. (1974). “Raoul de Presle”, in *Histoire littéraire de la France 40*, 113-186.
- DMF2009 = ATILF (CNRS & Nancy-Université), 2009. *Dictionnaire du Moyen Français (DMF2009)*, Nancy, ATILF/CNRS & Nancy-Université, site internet : <http://www.atilf.fr/dmf>.
- DUVAL, F. (2007). *Lectures françaises de la fin du Moyen Âge. Petite anthologie commentée de succès littéraire*, Genève : Droz.
- LABORDE, A. de (1909). *Les manuscrits à peintures de La Cité de Dieu de Saint Augustin*, Paris : Rahir.
- SOUVAY, G., PIERREL, J.-M. (2009). « LGeRM : Lemmatisation des mots en moyen français », *Traitement Automatique des Langues, volume 50* : 2, 149-172.